Information Theory
First Midterm
October 15, 2019

1) State the theorem called Jensen's inequality.

2) Give the definition of conditional entropy.

3) Let the random variable Y take values from the set $\{1, 2 \ldots, 6\}$ with probabilities

$$P(Y = 1) = \frac{1}{2},\ P(Y = 2) = \frac{1}{4},\ P(Y = 3) = \frac{1}{8},$$
$$P(Y = 4) = \frac{7}{64},\ P(Y = 5) = P(Y = 6) = \frac{1}{128}.$$

Construct the binary Shannon-Fano code for this distribution and decide whether it has optimal average length among the prefix codes encoding the value of Y.

4) We toss a fair coin several times until we will have two consecutive tosses with the same result or we already had 7 tosses. (That is, we stop after the first occasion of two consecutive heads or two consecutive tails or after having tossed the coin seven times.) Let X denote the random variable whose value is the number of tosses we make. Give an optimal average length binary encoding of X.

5) We choose two positive integers according to the uniform distribution from the sets

$$\{1, 5, 11, 23\} \text{ and } \{1, 7, 32, 64\},$$

respectively. Let U and V denote the two random variables whose values are the two randomly chosen numbers and let W and Z be their sum and product, respectively, that is,

$$W = U + V \text{ and } Z = U \cdot V.$$

Calculate the entropy values $H(W), H(Z), H(W|Z)$, and $H(Z|W)$.

6) Let X, Y, Z be three random variables, each taking its values on the set $\{0, 1\}$. We know that $H(X) = H(Y) = 1$ and $H(Z|X) = 1, H(Z|X, Y) = 0$. What are the smallest and the largest possible values the entropies $H(Z|Y)$ and $H(X, Y, Z)$ can take under these conditions?

Sketches of solutions for the exercises

3. Following the algorithm we learnt we find that the codewords for the Shannon-Fano code are:

$$0; 10; 110; 1110; 1111110; 1111111.$$

It is clear that this cannot have optimal average length, since we can simply shorten the last two codewords and simply obtain a prefix code:

$$0; 10; 110; 1110; 11110; 11111.$$

It is obvious that the latter has smaller average length.

4. The probability that the second toss is the same as the first one is $\frac{1}{2}$. The probability that the second toss is different from the first one but the third one is identical to the second is $\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$. Similarly, having the first similar than the previous one result at the ith tossing is $\frac{1}{2^{i-1}}$. This gives the probabilities for $X = 2, 3, 4, 5, 6$. The probability of $X = 7$ is the total remaining value: $1 - \sum_{i=2}^{6} \frac{1}{2^{i-1}} = \frac{1}{32}$. Thus the distribution for X is

$$(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{32}, \frac{1}{32}).$$

Constructing the Huffman code for this distribution we obtain the code

$$0; 10; 110; 1110; 11110; 11111.$$

5. One can easily see that all the possible products we can obtain as values of Z are different. (This is easiest to see by realizing that all numbers not equal to 1 are distinct primes plus two different powers of 2.) So the product will determine what were the numbers we multiplied and thus it will also determine their sum. Therefore Z determines W thus $H(W|Z) = 0$. The number of possible products is 16 and each has the same probability, thus $H(Z) = \log_2 16 = 4$. Among the possible sums there are only two equal ones: $1 + 11 = 12 = 5 + 7$, all other sums are different. Thus the sum being 12 has probability $2 \cdot \frac{1}{16}$, while the other 14 values have probability $\frac{1}{16}$ each. This gives the entropy value $H(W) = \frac{14}{16} \log_2 16 + \frac{1}{8} \log_2 8 = \frac{31}{8}$. Finally, by $H(Z|W) + H(W) = H(W|Z) + H(Z)$ we obtain that $H(Z|W) = 0 + H(Z) - H(W) = \frac{1}{8}$.

6. Using the Chain rule we have

$$H(X, Y, Z) = H(X) + H(Z|X) + H(Y|X, Z) \geq H(X) + H(Z|X) = 2.$$

On the other hand $H(X, Y, Z) = H(X) + H(Y|X) + H(Z|X, Y)$. Since $H(Z|X, Y) = 0$, this implies

$$H(X, Y, Z) = H(X) + H(Y|X) \leq H(X) + H(Y) = 2.$$

Thus we have $2 \leq H(X, Y, Z) \leq 2$, so

$$H(X, Y, Z) = 2.$$

The value of $H(Z|Y)$ is not determined, but we know $0 \leq H(Z|Y)$ by the nonnegativity of entropies and also $H(Z|Y) = H(Z, Y) - H(Y) = H(Z, Y) - 1 \leq H(X, Y, Z) - 1 = 2 - 1 = 1$. So we have

$$0 \leq H(Z|Y) \leq 1,$$

and both extremes can be attained: we can simply have $Z = Y$ in which case $H(Z|Y) = 0$, or we can have $Z = X + Y \pmod 2$, in which case $H(Z|Y) = 1$.